Computer-Aided Design and Drafting (CADD)
and
Geographic Information System (GIS)

Concepts and Terminology

The CADD/GIS Technology Center
 For facilities, infrastructure, and environment
Information Technology Laboratory
U.S. Army Engineer Research and Development Center
Vicksburg, Mississippi 39180

Concepts and Terminology

Computer-Aided Design and Drafting (CADD) Software.

Computer-aided drafting (CAD) technology was first introduced in the mid-1960's as a tool for the production of drawings without the use of traditional drafting tools. The drawings were created and displayed by manipulating graphic elements on the computer screen instead of drawing them by hand. Engineering design capability was added to many of the CAD programs and thus the name was changed to Computer-Aided Design and Drafting (CADD).

CADD technology has become the preferred method for the preparation, distribution, storage, and maintenance of architectural and engineering type drawings. With CADD systems, graphic three-dimensional or two-dimensional digital data are placed on various drawing layers that can be selectively displayed and edited.

Basic CADD Software. The two predominant commercially available CADD software packages currently used by DoD and other organizations are AutoCAD (AutoDesk) and MicroStation (Bentley). MicroStation and AutoCAD are "basic" CADD products in that they provide fundamental drafting and design capabilities.

CADD Application Software. CADD application software packages operate "on-top-of," or in conjunction with, the basic CADD software package (e.g., MicroStation and AutoCAD).

Advanced application packages assist in the performance of specialized design or analytical functions such as highway design, site design, architectural design, and survey/mapping. Some products store attribute data with existing graphic elements. Others provide linkages between graphic elements and data in external files or an external database. The advanced CADD software application packages, customer support libraries, user commands, and menus may be written for a specific hardware/software platform.

Extension application packages enhance drafting and modeling productivity or some other specific function.

Facility Management (FM) Software.

FM software can either: (1) Consist of a "nongraphic" software program utilizing a database, or (2) Consist of a software program based on CADD or GIS technology (with integrated database) with an emphasis placed on providing enhanced capability to manage utility systems and facilities data. A CADD/GIS based FM system permits the linkage of digital maps with a database management

system (DBMS) in a manner which permits the manipulation and query of the type of information that is needed for the management of facilities.

Geographic Information System (GIS) Software.

GIS technology provides a computerized mechanism for capturing, verifying, storing, manipulating, querying, analyzing, and displaying geospatial data referenced to their location on earth. GIS technology has evolved from the computerized applications of cartography and map analysis which were first developed in the 1960's. However, its practicality and widespread use have only been realized in fairly recent years due to the rapid advancement of computer hardware and operating system software capabilities.

There are two primary graphic structures for GIS: raster and vector.

In raster GIS, graphic features are stored as an array of cells (pixels) referenced by row and column numbers.

In vector GIS, geographic features (i.e., entities) are represented by three primitive geometrical entities: points, lines, and polygons. Data associated with a geographic entity are stored in a database for subsequent access. Without an attached attribute database, a line is simply a string of x-y coordinate pairs (vector) or cells (raster), with a known location relative to some coordinate system. The attribute data for a line, stored within the database, provide additional information about the line (e.g., the type of entity represented by the line, its characteristics, etc. For example, a waterline, 8 in. in diameter, ductile iron pipe material, constructed in 1986 by ACME, Inc., etc.).

The basic components of a GIS are: (1) the software (including application software and DBMS software), (2) the computer platform (i.e., hardware and operating system), (3) the data, (4) the people who use, operate, and maintain the GIS, and (5) procedures. There are considerable differences in the capabilities of the commercially available GIS software on the market today. Until recently, with the development and proliferation of the Microsoft Windows NT operating system and the advancement in Personal Computer (PC) capabilities, most GIS software operated on UNIX-based workstations. Today, the GIS market is growing and changing at a rapid pace. Also, the gap between the capabilities and operational characteristics of traditional CADD and GIS software is rapidly narrowing.

GIS combines the use of CADD technology and Relational Database Management System (RDBMS) technology to relate data to features on digital maps and drawings and to allow for the creation, storage, maintenance, retrieval, query, analysis, and display of various geospatial information.

GIS differs from CADD, AM, and AM/FM in that spatial relationships among all data elements are defined through a convention known as topology. Topology provides a mechanism for describing the location, geometry, and characteristics of map features, as well as how linear map features are connected, how areas are bounded, and which areas are contiguous. A GIS uses a special database structure to define map topology. All map features are defined as either a point (node), line (arc), or polygon (area). GIS systems can also represent network topology similar to AM/FM systems.

Point features are used to represent map or drawing features which are identified as symbols or labels (i.e., objects whose shape or boundary is too small to be shown as a line or polygon at a particular map scale). Points at the beginning, end, and intersections of lines represent nodes (and are used primarily for utility networks). Incident locations are specific point features where an action occurred (e.g., where a break was repaired in a waterline). For most GIS software, each point is assigned a unique identification number and is located by a coordinate value (i.e., x, y, and sometimes z (elevation) geographic coordinate values).

Line features are described by a series of coordinate pairs (i.e., x, y, and sometimes z) and are assigned a unique identification number. Two coordinate pairs, one for the beginning and one for the end of the line, are needed to fully describe a straight line. Many coordinate pairs may be needed to fully describe a curvilinear feature. Line features are used to represent objects that are too narrow to display as a polygon at a particular map scale (e.g., a highway or stream). Line features are also used for an object that has no width (e.g., a contour line).

Polygon features represent a defined bounded area. A polygon's boundary is represented by a closed line or series of lines. Each polygon has a uniquely identified centroid, which is a point located anywhere within the area. A coordinate pair defines the location of the centroid. A polygon centroid has no correlation to the actual mathematical center of the polygon feature.

The true value of a GIS is that it permits the query and analysis of the graphic and associated attribute data. Attribute data are nongraphic data which describe or provide detailed information concerning spatially related features on a map or drawing. The attribute data are stored in a relational database management system (RDBMS) separate from the graphics data. Each attribute table is electronically linked, or attached, to the appropriate spatial feature to which it is related. The attribute table containing data describing a polygon feature is attached to the polygon's centroid, feature (i.e., boundary), or label. The attribute table containing data describing a point, polygon, or line is attached directly to the point, line, or polygon.

There are considerable differences in the capabilities of the commercially available GIS software on the market today. With the growing popularity of GIS technology, numerous software packages are being developed which are

advertised as providing GIS capability. For the purposes of this manual, the various GIS software packages on the market today will be broken down into the following two broad categories:

- Workstation GIS software
- Desktop GIS software

Workstation GIS Software. Workstation GIS software provides all of the data capture, storage, editing, retrieval, analysis, and display capability that is available from GIS technology. Because of its extensive capability, workstation GIS requires either a UNIX or Windows based workstation platform to operate.

The two predominant commercially available GIS software programs currently used by organizations are ARC/INFO (Environmental Systems Research Institute, Inc. (ESRI)) and Modular GIS Environment (MGE) (Intergraph). Both ARC/INFO and MGE/MGA are vector-based GISs. Application software is available to analyze and clarify raster data.

ARC/INFO is a stand-alone product that requires some programming skills to operate. ARC/INFO is currently available for the UNIX workstation platform. ESRI has recently developed a Windows version.

MGE is available for both the UNIX and Windows platforms and runs "on-top-of," or in conjunction with, MicroStation (Bentley, Inc.).

Desktop GIS Software. Desktop GIS software is capable of providing a significant amount of data capture, storage, editing, retrieval, analysis, and display capabilities, but not to the extent provided by workstation GIS software. Desktop GIS software is used extensively for AM/FM-type applications.

ARC/CAD (ESRI), GeoMedia (Intergraph), AutoCAD Map (Autodesk), GeoGraphics (Bentley), ArcView (ESRI), and MapInfo (MapInfo) are commercially available GIS products which fall into this category.

Database Management Systems (DBMS).

A database consists of a structured and organized collection of information. A DBMS is a computer program which provides for the management of the data, or information, contained in a database. A relational DBMS (RDBMS) is a computer program which provides a means of managing the related data contained in one or more database tables.

The computer language which has been developed for organizing, managing, interacting, and retrieving the data stored in a RDBMS is called Structured Query Language (SQL). SQL is an integral part of a RDBMS and provides the following functions:

1. *Data Definition* - The organization and structure of the data and the relationships between the data can be defined by the user.

2. *Data Retrieval* - Stored data can be retrieved from the database and used by the user or an application program.

3. *Data Manipulation* - The user of an application program can update the database by adding new data, removing old data, and modifying previously stored data.

4. *Access Control* - A user's ability to retrieve, add, and modify data can be restricted, thereby protecting stored data against unauthorized access.

5. *Data Sharing* - Data shared by concurrent users can be coordinated, thereby ensuring that they do not interfere with one another.

6. *Data Integrity* - Integrity constraints in the database can be defined, thereby protecting it from corruption due to inconsistent updates or system failures.

The American National Standards Institute (ANSI) and the International Standards Organization (ISO) published standards for SQL in 1986. The ANSI/ISO SQL standards were significally expanded in 1992. SQL is also included in the U.S. Federal Information Processing Standards (FIPS).

Oracle (Oracle Corporation) and Informix (Informix Software Inc.) have traditionally been the most widely used RDBMS. dBase (Borland) has also been widely used as a DBMS (i.e., non-SQL). INFO (Henco Software, Inc.) is a DBMS (i.e., non-SQL) which is bundled and distributed with ARC/INFO (ESRI). The growing popularity of Windows NT (Microsoft) has resulted in the growth in popularity of RDBMSs complying with Microsoft's Open Database Connectivity (ODBC) standard, such as Access (Microsoft). Other popular RDBMSs include Foxpro (Microsoft), Paradox (Borland), and SQL Server (Microsoft).

CADD & GIS Application Software.

CADD and GIS Application software packages can operate as "stand-alone," or in conjunction with the CADD or GIS software to add some additional specialized capability or function (e.g., performance of specialized data conversion, map generation, scanning, or viewing/query/analysis functions).

Graphic Concepts.

Vector graphics. Vector elements are graphical objects that have a precise direction, length, and shape. In CADD software, the vector graphical objects can be points, lines, polygons, arcs, rectangles, circles, splines, text, ellipses, elliptical arcs, arc wedges, elliptical arc wedges, and symbols. In GIS software, the vector graphical objects can be points, lines, or polygons.

Grouped vector objects are either in the form of graphic groups or complex elements such as cells or blocks. Vector graphics are particularly well suited for processes where map development and modification are heavy and true two- and three-dimensional accuracy is required. Vector graphic images can exist in two- or three-dimensional design environments. They are created using CADD or GIS software packages. In addition, the ability to deal with geometric elements as unique entities provides a powerful linkage for images to be grouped with nongraphical data attributes. Output from CADD/GIS software packages creates a "proprietary" binary data file that may have limited portability across various platforms and vendor packages, unless it is converted into a "neutral" data format.

Two-dimensional vector entities are represented by x- and y-coordinate points, taking the form of a single point, string of points (e.g., lines, arcs, and splines), and closed lines (e.g., polygons, rectangles, circles, and ellipses).

Three-dimensional vector entities are represented by x-, y-, and z-coordinate points.

Raster Graphics. Raster graphics, or bit-mapped graphics, are digital images stored as arrays of pixels for display and modification. In raster data there are no lines, circles, or polygons, only pixels that are grouped to give the appearance of these elements.

In a monochrome (black and white) graphic system each pixel corresponds to one binary digit (bit) on the display screen. A bit is the smallest unit of information that a computer can handle. Color graphic systems require more bits per pixel to create colors. The need for more memory and memory management increases as both color and screen resolution increase.

Raster graphics are commonly created in one of two ways:

a. Use of "paint" type computer graphics programs (e.g., desktop publishing, image editing, and paint) to draw and edit images.

b. Quick capture by scanning the image of existing hard-copy maps and converting them to a digital format (bit-mapped). Once captured digitally, the maps can be archived, distributed, or converted to a vector format. The

scanning can be accomplished by either electronic scanning equipment or digital cameras.

Some of the more common raster (bit-mapped) file formats in use include:

a. Microsoft Windows Paintbrush (PCX).

b. Tag Image File Format (TIFF).

c. Graphics Interchange Format (GIF).

d. Bit-Mapped Picture (BMP).

e. Intergraph I/RASC

The type(s) and versions(s) of raster graphics software packages used by the target organization should be addressed in the contract technical provisions.

"Neutral" file exchange formats. CADD graphic files converted to "neutral" file exchange formats such as Drawing Exchange Format (DXF) and Initial Graphics Exchange Specification (IGES) can be converted by most CADD and GDS software to their native file formats. However, caution should be exercised because the "neutral" file exchange formats may not transfer all the graphic elements (e.g., arcs, splines) stored in the native CADD file format. Basic CADD software packages and add-on or third-party CADD software application packages are typically employed to develop CADD drawings. Few of the "neutral" file exchange formats currently available have reliable mechanisms to transfer the wide variety of electronic linkages between graphic elements and nongraphic attributes used in basic and advanced CADD software packages.

DXF, although widely supported and well known, was developed by AutoDesk to support the graphic elements generated by AutoCAD. Because DXF maps AutoCAD graphic entitities, it is only "neutral" to the extent that AutoCAD entities are common to most other CADD software packages.

The set of graphic elements available in IGES is sufficient to represent the graphics from almost any CADD software package. IGES is a public standard.

Both DXF and IGES file exchange formats will exchange two-dimensional drawings, two- and three-dimensional wireframe models, and simple three-dimensional surfaced solids. "Neutral" file exchange formats, such as DXF and IGES, were established for the exchange of CADD-generated graphic entities. They can tell you what in the world it is with a fair amount of accuracy but it does not tell you where in the world it is located.

Many problems can be prevented by: (1) Avoiding the use of certain graphic elements; (2) Converting complex graphic elements to simpler forms; (3) Using simple and standard text fonts; and (4) Careful control of layering and symbology.

When translating to and from "neutral" file exchange formats, each CADD package tries to transfer the appearance of the "neutral" file exchange format's graphic elements into the types of graphic elements it supports. This works well for simple graphic elements (e.g., lines, circles, and rectangles), but does not work well for complex graphic elements. The "neutral" file exchange format's graphic elements not supported by the CADD package may be lost or modified when the "neutral" file is imported.

The electronic linkage between graphic and nongraphic elements may be lost when converting a file to a "neutral" file exchange format. Different CADD packages have different ways of storing nongraphic data. Some CADD packages can store a small amount of information with the graphic element in an application-defined format or, alternatively, have one or more electronic linkages to external databases where a larger amount of data can be stored.

The Spatial Data Transfer Standard (SDTS), also Federal Information Processing Standards (FIPS) 173, was approved by the National Institute of Standards and Technology (NIST) in 1992, for the purpose of providing a "neutral" format for the exchange of graphic and nongraphic geospatial and related data. The SDTS provides an exchange mechanism for the transfer of spatial data between dissimilar computer systems. The SDTS specifies exchange constraints, addressing formats, structure, and content for spatially referenced vector and raster data. The SDTS was approved as FIPS 173 in July 1992. The FIPS 173 implementation became effective February 15, 1993; use of FIPS 173 as the spatial data transfer mechanism was mandatory for Federal agencies on February 15, 1994. It is also available for use by state and local governments, the private sector, and research and academic organizations.

CADD and GIS Integration Considerations

A decision should be made whether or not survey, mapping, or digital geospatial data will later be used for both CADD and GIS applications. Both technologies can be used to create maps that have a similar outward appearance.

CADD technology treats digital data as electronic drawings that are basically made up of graphic elements organized into "layers" or "levels." CADD technology has become more sophisticated with the development of the capability to store basic nongraphic data about graphic elements in external databases. The stored data can be queried for a variety of purposes including design analysis or facilities management.

GIS technology is more complex because it must accurately store both graphic (map or drawing) and nongraphic (database or attribute) data for analysis and display. GIS technology can be used to simulate extremely complex real-world events and situations.

At most organizations, digital survey maps and drawings have been prepared using CADD technology. Today, with the growing popularity of GIS technology, the life cycle use of the electronic data should be evaluated in determining the most useful and efficient means for data acquisition and digital map development. Many organizations use CADD technology for the development of digital survey maps whose primary use is the development of design (architectural and engineering) drawings for construction-type projects. GIS technology should be used for the development of maps created specifically for planning, design, operations and maintenance, facility management, and disposal functions where the analysis of stored nongraphic data is the primary concern.

When CADD-generated data files are used by a GIS, the following guidelines in data structure should be followed:

a. The edges of all digitized maps must exactly match digitally with those of all adjacent maps.

b. The digital representation of the common boundaries for all graphic features must be exactly the same, regardless of level/layer.

c. Lines and line strings which represent the same graphic element must be continuous (i.e., not broken or segmented), unless that segmentation reflects a specific visual line type. Lines/strings representing the same type of data must not cross except at shared vertices.

d. Polygons must be closed (i.e., the first x- and y-coordinates must exactly match the last x- and y-coordinates). Each polygon must have a single unique centroid to which attributes (i.e., an attribute table) can be attached. Polygons of the same coverage must not overlap and must cover the area of interest completely (i.e., have no gaps in coverage).

e. All graphic elements that connect must connect digitally, without overlaps or gaps.

f. Straight lines must be represented by only the beginning and ending x- and y-coordinate points. Line strings must not cross back on themselves or be of zero length (i.e., points).

National Geospatial Standards

The increased use of GIS within the local, state, and federal gove rnment, business activities, academic and international communities in recent years has resulted in an increased demand for geospatial data. The term "geospatial data" refers to any information that identifies the geographic location and characteristics of natural or constructed features and boundaries on the earth.

The goal of the Federal Geographic Data Committee (FGDC) is to create and adopt standards that promote the coordinated development, use, sharing, and dissemination of geospatial data on a national basis. The FGDC is an interagency committee established by the 1990 revision of Office of Management and Budget (OMB) Circular A-16, entitled "Coordination of Surveying, Mapping, and Related Spatial Data Activities". The Department of Interior was designated as the lead agency for the FGDC. OMB Circular A-16 also established a process to promote the development of a national spatial framework for an information-based society with the participation of Federal, state, and local governments, and the private sector, and to reduce the duplication of effort.

Executive Order 12906, "Coordinating Data Acquisition and Access: The National Spatial Data Infrastructure" (NSDI), which was signed by President Clinton on 11 April 1994, requires that all Federal agencies use the FGDC Metadata Standard to document new geospatial data and make them electronically accessible through the use of a National Geospatial Data Clearinghouse. Executive Order 12906 also assigned authority for the development of na tional geospatial data standards to the FGDC. The NSDI Clearinghouse is intended to be a distributed, electronically connected network of geospatial data producers, managers, and users. When fully functional, the NSDI Clearinghouse will allow its users to electronically (via the Internet) determine what geospatial data exists, find the data they need, evaluate the usefulness of the data for their applications, and obtain or order the data as economically as possible.

The NSDI will provide a base or structure of relationships among data producers and users that will facilitate data sharing. The FGDC comprises 14-plus subcommittees and working groups whose focus is on developing geospatial standards. The FGDC program ensures that standards are created under an open consensus, with participation by non-federal and federal communities and that all standards from the FGDC subcommittees and working groups are to be integrated. More information about the FGDC and individual standards development can be found through the FGDC Internet Homepage at URL: http://www.fgdc.gov . The FGDC data standards are being incorporated into the SDS/FMS as they are approved. The FGDC data standards for Vegetation, Wetlands, and Soils were incorporated into the TSSDS/TSFMS Release 1.80.

Geospatial Metadata

The ability to share existing geospatial data between organizations is essential. This capability is dependent upon a valid data schema (capability to locate data), a data dictionary (capability to understand the data), and the data quality. Metadata or "data about data" are a key to developing this capability. Metadata describe the content, quality, condition, and other characteristics of data.

The FGDC approved the standard on 8 June 1994. The FGDC sponsored a public review and test of a draft spatial metadata content standard. The draft standards proposed a set of data elements that identified the data, described the projection, status, data content lineage and processing, and quality of the data. It also identified the custodians and access conditions for the data. Comments were sought on the content, completeness, and usability of the standard, and the identification of elements which should be mandatory for geospatial data cataloging and exchange. Executive Order 12906, Coordinating Data Acquisition and Access: The National Spatial Data Infrastructure, which was signed on 11 April 1995, requires all Federal agencies use the Metadata Standard to document their geospatial data.

Provisions should be included in the technical specifications requiring a Architect-Engineer contractor to prepare and deliver the metadata files along with the final contract deliverables. Internet sources of detailed information concerning the content and format for the metadata files can be obtained from the following URL addresses:

FGDC Internet Homepage - http://www.fgdc.gov

FGDC Clearinghouse - http://www.fgdc.gov/clearinghouse

FGDC Metadata - http://www.fgdc.gov/metadata

USACE National Geospatial Data Clearinghouse Node -
   http://corps_geo1.usace.army/mil

Metadata Generation Software has been developed to assist in the preparation of the required metadata files. Some sources of Metadata Generation Software are:

- USACE National Geospatial Clearinghouse Node - CORPSMET is available for download at no cost. CORPSMET is a DOS application. A Windows 95/NT version of CORPSMET, as well as a HYPACK metadata generation tool, has been developed (CORPSMET95).

- National Oceanic and Atmospheric Administration (NOAA) - Metadata Generation Software is available for download at no cost from ftp://www.fgdc.gov/pub/tools/NOAA.

- National Biological Service - Meta Maker (a Microsoft Access application) is available for download at no cost from http://www.emtc.nbs.gov/http_data/emtc_spatial_applications.

- ARC/INFO (ESRI) has Document.aml for metadata generation.